

TRATAMIENTO DE LOS SENTIDOS VERBALES PARA RECURSOS COMPUTACIONALES DE ENSEÑANZA DE LENGUAS

Morantè, Roser
morante@lingua.fil.ub.es
Martí, M. Antònia
amarti@lingua.fil.ub.es
Díaz, Lourdes
lourdes.diaz@trad.upf.es

Resumen

En esta comunicación se presenta una propuesta de tratamiento de los sentidos verbales para relacionarlos con niveles de aprendizaje. El estudio se enmarca en un proyecto explotación de la base de conocimiento léxico EuroWordNet como herramienta de apoyo al aprendizaje de segundas lenguas. Para definir los aspectos del conocimiento verbal que se representarán en EWN se adapta el modelo de análisis de predicados Pirápides definido para sistemas de PLN. La información sobre los usos verbales se extrae de diccionarios y corpus. Para determinar la gradación de los usos verbales por niveles de aprendizaje se toma como referencia el concepto de prototipo. Se parte de la hipótesis de que el parámetro de prototipicidad está directamente relacionado con el nivel de aprendizaje: los usos más prototípicos de los diversos aspectos del conocimiento verbal se enseñarán en niveles más básicos. En este trabajo presentamos los parámetros que se han definido y su aplicación al verbo *llevar*.

1 Introducción

El marco general de este trabajo es la enseñanza de segundas lenguas asistida por ordenador, en concreto, la explotación de la base de conocimiento léxico EuroWordNet [Vossen 98] como herramienta de apoyo al aprendizaje del uso verbal. En [Morante & Martí 98] se exponen las líneas generales de la realización del proyecto, cuyo objetivo es la integración de EWN en un contexto didáctico, que permita al usuario disponer de

información relevante y sistematizada en su proceso de reestructuración de la L2 [Liceras 96].

El diseño de herramientas para asistir el aprendizaje de lenguas requiere la tarea previa de definir la información que es conveniente presentar en cada nivel de aprendizaje [Self 92]. Para relacionar los sentidos verbales con niveles de aprendizaje hemos establecido diversos parámetros, que miden el nivel de prototipicidad de los diferentes usos y sentidos del verbo. Partimos, pues, de la hipótesis de que el grado de prototipicidad está directamente relacionado con el nivel de aprendizaje. En esta comunicación se expondrán los parámetros establecidos y se aplicarán a algunos de los usos del verbo *llevar*. Cabe señalar que paralelamente se están realizando experimentos para validar la hipótesis de partida y que, por tanto, presentamos un trabajo en curso.

En la sección 2 se presentará la metodología de análisis de los sentidos verbales. En la sección 3 se presenta el concepto de prototipicidad y los criterios para aplicarlo a los usos verbales. En la sección 4 se aplican a algunos de los usos verbo *llevar* y, finalmente, en la sección 5 se presenta el trabajo en curso y las líneas futuras.

2 Metodología de análisis de predicados

Una de las tareas previas al diseño de la explotación de EWN como herramienta de apoyo al aprendizaje es la definición de la información verbal de que el usuario debe disponer. Para ello se ha adaptado la propuesta del modelo Pirápides de análisis de predicados [Fernández *et al.* 98], definido inicialmente para sistemas de PLN. Consideramos que los

resultados de investigaciones en lingüística computacional son integrables en el diseño de herramientas informáticas de aprendizaje de lenguas, ya que tanto los programas de PLN como el material de aprendizaje de lenguas deben explicitar el conocimiento implícito de los usuarios.

Siguiendo la hipótesis de [Levin 93], el mencionado modelo asigna a cada verbo componentes de significado y alternancias de diátesis por el hecho de pertenecer a una clase semántica. Los componentes de significado son una reformulación de las primitivas propuestas por [Talmy 89]. Cada lengua presenta esquemas particulares de lexicalización de los componentes. Por tanto, en función de las diferencias entre la L1 y la L2 del estudiante se deberá hacer énfasis en ciertos modelos de lexicalización y alternancias de diátesis específicos. En cuanto

a las alternancias de diátesis, éstas se conciben como recursos gramaticales característicos de una lengua o un grupo de lenguas usados para expresar oposiciones semánticas generales. Remitimos a [Fernández *et al.* 98, Vázquez 98, y Vázquez *et al.* 99] para una descripción del mencionado modelo, cuyos fundamentos daremos por supuestos en este trabajo.

La información obtenida para cada sentido verbal mediante este estudio queda reflejada en la tabla 1. La información relativa a las características idiosincráticas de cada sentido verbal se extrae de una fuente de información auxiliar, el corpus de producciones escritas de nativos, LEXESP. En la enseñanza de segundas lenguas es básico que el estudiante disponga de materiales auténticos [Little 98]. Por esta razón, consideramos que el material didáctico de aprendizaje de léxico debe estar basado en datos extraídos de corpus.

Tabla 1. Información por sentidos verbales

VERBO: <i>LLevar</i>	SENTIDO: 1
• Glosa esquematizada	Alguien [1] lleva algo/a alguien [2] a algún sitio/a alguien [3]
• Clase semántica del predicado	Transferencia
• Componentes de significado	Iniciador [1], entidad [2], trayectoria-meta [3]
• Información sobre los componentes de significado:	
• Categoría y función sintáctica referida a la variante de la alternancia indicada en la glosa. Para los SP se indica también la preposición entre paréntesis:	[1] NP/SUBJ, [2] NP/DOBJ, [3] PP (a)/CPREP
• Restricciones de selección:	[1] animado, [2] móvil, [3] lugar
• Cadenas léxicas prototípicas que realizan el componente.	LLevar a alguien a la ruina Llevar a alguien de cabeza
• Alternancias sintácticas. Cada alternancia se ejemplificará con ejemplos extraídos del corpus.	Alternancia de trayectoria, pasiva, impersonal, dativo de interés
• Información sobre otras piezas léxicas con las que establece relaciones semánticas en términos de EWN: hipónimos, hiperónimos, sinónimos, antónimos, involved, role.	
• Bateria de ejemplos extraídos del corpus LEXESP.	

La metodología empleada para el análisis de predicados se basa en el modelo Pirápidés. Los pasos que se siguen son los siguientes:

- Extracción de ocurrencias.
- Detección de sentidos tomando como guía los sentidos establecidos en diccionarios y añadiendo sentidos mediante un criterio laxo de sinonimia.

- Análisis semántico, de componentes de significado y restricciones de selección.
- Análisis sintáctico, forma de realización de los componentes, alternancias de diátesis.
- Análisis léxico centrado en el estudio de las coocurrencias léxicas.

Una vez el verbo ha sido tratado en todos

estos aspectos, se le aplica la escala de prototipicidad para determinar qué usos se corresponden con qué niveles. En el apartado siguiente se explica en qué consiste la escala.

3 Criterios para establecer la prototipicidad de los sentidos verbales

En el apartado anterior se ha expuesto el tipo de información que se incluirá para cada sentido verbal. Una vez EWN esté enriquecido será necesario diseñar la forma de acceso a la información que contenga. Para que la consulta a la herramienta sea efectiva, es imprescindible que al alumno se le presente la información adecuada para su nivel de competencia. Por tanto, el conocimiento verbal deberá organizarse siguiendo algún tipo de criterio, como podría ser el basado en syllabus.

Sin embargo, ya se ha avanzado, que la asignación de sentidos y usos verbales a niveles de aprendizaje se realiza tomando como referencia el concepto de prototipo, desarrollado en el marco de la lingüística cognitiva, entre otros, por [Kleiber, Lakoff 87, Langacker 87, Rosch 77, Taylor 89]. Esta teoría propone un concepto de categoría que se opone a la concepción aristotélica, según la cual cada categoría tiene asociados un conjunto de condiciones necesarias y suficientes que deben cumplir todos los elementos que la conforman.

Existen diferentes enfoques sobre la definición de prototipo. [Coleman & Kay 81] proponen un modelo de definición de prototipos basado en un conjunto de rasgos binarios, de manera que los elementos prototípicos son aquellos que se definen mediante todos los rasgos, mientras que los no prototípicos tienen un grado de pertenencia a la categoría proporcional a los rasgos que los definen. [Kronenfeld 88] y [Taylor 89] consideran que un prototipo es el mejor ejemplar de la categoría y constituye su núcleo.

Partiendo de lo señalado por algunos autores [Coleman & Kay 81], en nuestro trabajo consideraremos la hipótesis de que el parámetro de prototipicidad está directamente relacionado con el nivel de aprendizaje: los usos más prototípicos de los diversos aspectos del conocimiento verbal se adquirirán antes

que los periféricos. Así, definiremos el uso prototípico basándonos en tres fuentes distintas:

- Las unidades de análisis que caracterizan el modelo de entrada verbal del proyecto Pirápidas: componentes de significado, alternancias de diátesis y estructura eventual.

- La definición de cadena causal y evento ideal de [Croft 90]. En este modelo un verbo se define por denotar un segmento de la cadena causal [causa – proceso – estado], que es un modelo cognitivo idealizado (MCI). Este MCI permite representar de forma unificada los tres tipos aspectuales básicos de eventos, que, según Croft, se relacionarán prototípicamente con ciertas clases semánticas. Cuando un evento se expresa desde una perspectiva que no es la prototípica, la perspectiva será marcada. Croft define como prototípicos los eventos que se ajustan al MCI y como periféricos los que no se ajustan. Esta diferenciación se manifiesta en que los eventos ideales se expresan gramaticalmente de la misma manera en todas las lenguas, mientras que los no ideales presentan variaciones.

evento	segmento de la cadena
causativo	causa-proceso-estado alguien rompió la ventana
incoativo	proceso-estado la ventana se rompió
estativo	estado la ventana está rota

- La jerarquía temática, según la concepción de Bresnan & Kanerva [Bresnan & Kanerva 92]. Se trata de una jerarquía de prominencia de papeles temáticos. Aunque en nuestro modelo de análisis de predicados no tomamos los roles temáticos como unidades, la jerarquía puede aplicarse también a los componentes de significado. El grado de prominencia se relaciona con la función sintáctica. Así, en posición de sujeto aparecerá prototípicamente el argumento que tenga papel *agente*.

Jerarquía temática:

agente > beneficiario > meta/experimentador >
instrumento > paciente / tema > locativo

Equivalencia con componentes de significado:

Agente: iniciador.
Paciente, tema : entidad.
Beneficiario, goal: meta.
Instrumento: instrumento.
Experimentador: experimentador.
Locativo: locativo.

A continuación exponemos los factores que pueden intervenir en el grado de prototipicidad, cuya validez se comprobará de forma experimental. En la sección 4 estos criterios se aplicarán a ejemplos del verbo *llevar*.

3.1 Clase semántica y tipo eventual

Tomando como referencia la hipótesis de Levin, el trabajo del proyecto Pirápidas y la teoría de Croft, partimos del supuesto de que un verbo pertenece prototípicamente a una clase semántica y a un tipo eventual. En el caso del verbo *llevar* la clase es la de transferencia [Morante *et al.* 98] y el tipo eventual el causativo [Croft 90]. No obstante, como se puede comprobar en los ejemplos del corpus, el verbo *llevar* no aparece en todos los sentidos ni como un verbo de transferencia, ni como causativo. Consideraremos prototípico el uso ideal del verbo y periféricos los demás.

3.2 Componentes de significado

Cada sentido del verbo, en función de la clase semántica a la que se asigna, recibe unos componentes de significado. Por ejemplo, el verbo *llevar*, como verbo de transferencia, recibe los componentes básicos iniciador, entidad y trayectoria. Estos componentes se realizan mediante unos recursos gramaticales (expresión sintagmática, lexicalización, incorporación, gramaticalización) y, en el caso de la expresión sintagmática, mediante unas funciones sintácticas en las estructuras no marcadas.

En referencia a estos componentes de significado, la prototipicidad se valora, para cada sentido, en función de:

- Nivel al que pertenecen los componentes del verbo, según los niveles establecidos en el proyecto Pirápidas. Los sentidos con componentes especializados (p.e. manera en *zarandear*) son menos prototípicos, por ser más especializados.

- Forma como se realizan los componentes: la expresión sintagmática es más prototípica

que la lexicalización (transparencia morfológica, *martillo-martillear*), y la lexicalización más que la gramaticalización (incorporación en el verbo del componente de significado sin transparencia morfológica).

- Relación de la posición del componente en la jerarquía temática respecto de la función que ocupa en la estructura no marcada del sentido verbal. Por ejemplo, en el sentido de transferencia (*alguien lleva algo a algún sitio*), el iniciador ocupa la posición de sujeto y, por tanto, nos encontramos ante una proyección prototípica. En cambio, el *llevar* de experimentación realiza un paciente en función de sujeto (*alguien lleva mucho tiempo en un estado*) y, por tanto, se tratará de un uso menos prototípico del verbo.

- Tipo de entidad en que se proyecta el componente respecto de las restricciones de selección indicadas en la estructura argumental. Los usos del verbo que transgredan en algún grado las restricciones de selección serán menos prototípicos. Por ejemplo, el *llevar* de transferencia requiere un iniciador animado. Sin embargo, en los usos metafóricos puede aparecer un iniciador causativo abstracto (*su egoísmo lo llevó a la ruina*). Por lo que respecta a las restricciones de selección, también se considerarán menos prototípicos los verbos que impongan restricciones de selección más especializadas (*importar* será menos prototípico que *llevar*).

3.3 Productividad

Algunos de los sentidos del verbo puede ser asignados fácilmente a una clase semántica (*llevar algo a algún sitio*, transferencia) y son muy productivos, puesto que el verbo puede combinarse con muchas piezas léxicas diferentes. Otros sentidos son relativamente transparentes y conservan cierta productividad (p.e.: *llevar algo en las venas*, *llevar algo en el interior*, *llevar algo en la sangre*, *llevar algo muy dentro*; *llevar bigote*, *barba*, *el pelo rizado*, etc.). En estos se pierde parte del concepto de transferencia, pues lo transferido es una propiedad inalienable, un atributo del sujeto. Finalmente, otros sentidos han perdido toda transparencia (*llevarse a alguien de calle*, *llevar algo a cabo*, *llevarlo claro*). Cuanto más productivo, sea un sentido, más prototípico será.

3.4 Definición del uso prototípico

A partir de los factores expuestos anteriormente se define el uso prototípico del verbo:

- La clase semántica del sentido es la prototípica del tipo aspectual.
- Expresa los componentes básicos de la clase semántica a la que pertenece prototípicamente. Se considerarán extensiones los usos del verbo que pertenecen a otra clase semántica.
- Todos los componentes de significado se expresan sintagmáticamente.
- El significado del verbo es transparente.
- Los componentes de significado se expresan en la función sintáctica que les corresponde en función de la jerarquía temática.
- Las entidades en que se realizan los componentes se ajustan a las restricciones de selección marcadas en la estructura argumental. Las violaciones de las restricciones darán lugar a usos metafóricos.

4 Aplicación de la metodología: gradación por niveles de los usos del verbo llevar

En esta sección aplicaremos los criterios arriba expuestos para clasificar los usos del verbo *llevar* por niveles de prototipicidad. Estos criterios serán la base de partida para relacionar conocimiento verbal y niveles de aprendizaje. No obstante, habrá que tener en cuenta también otros factores. Presentamos sólo algunos de los usos del verbo agrupados por sentidos, de acuerdo con el criterio de sinonimia ya mencionado. Para cada sentido indicamos los componentes de significado y los identificamos en el ejemplo mediante un índice.

Los ejemplos aparecen ordenados por orden de más a menos prototipicidad para cada grupo, cuya ordenación se ha realizado en función de dos criterios, las restricciones de selección y la transparencia de significado. Se sitúan primero los ejemplos en los que los componentes se realizan mediante entidades no abstractas y en los que el significado es transparente.

A su vez, los grupos también se presentan por orden de mayor a menor prototipicidad. El

grupo A) corresponde al significado básico de *llevar* como verbo de transferencia con todos los componentes expresados de forma sintagmática. En los grupos B), C), D) y E) no se expresa el componente trayectoria. Queda inespecificado en todos los grupos, aunque en los ejemplos de los grupos D) y E) el verbo aparece en su forma pronominal. Por esta razón se consideran menos prototípicos que B) y C). B) se ha considerado más prototípico que C) y D) que E) por el hecho de que en B) y D) el iniciador es agentivo, mientras que en C) y E) es experimentador. Finalmente en el grupo F) se incluye un sentido del verbo lexicalizado. Por supuesto, esta ordenación deberá comprobarse experimentalmente.

A) TRANSPORTAR: [1] iniciador, sujeto [2] entidad, objeto [3] trayectoria, complemento

1. Me[2] llevaron a una casa en ruinas[3] y había un colchón.
2. Llevaron la lengua de Castilla [2] al corazón del Africa Negra[3].
3. ... los escándalos [1] que llevaron a la cárcel[3] a algunos de los protagonistas del equipo [2]
4. ... la angostura mental [1] que llevo a la rigidez [3] y puede culminar en el extravío.
5. Cándido [1] llevaba en sus venas [3] el gusanillo de la aventura [2]

B) DIRIGIR: [1] iniciador, sujeto [2] entidad, objeto

1. El general [1] llevaba el mando de la operación [2]
2. La dirección [1] siempre llevaba la voz cantante [2]

C) TENER: [1] iniciador, sujeto [2] entidad, objeto

1. Hierro y Salinas [1] no llevar bigote [2] ...
2. No basta con llevar un apellido ilustre [2] ...
3. ... la camisa blanca [1] que llevo zurcidas en hilo las iniciales [2]

D) GANAR: [1] iniciador, sujeto [2] entidad, objeto [3] trayectoria, gramaticalizada

1. Rominger [1] se [3] llevó la etapa más bonita [2]
2. Juntos se [3] llevaron de calle dos títulos de liga [2]
- 3 La temporada trágica [1] se [3] llevó la palma [2]

E) EXPERIMENTAR: [1] iniciador, sujeto [2] entidad, objeto [3] trayectoria, gramaticalizada

1. Bertolucci [1] ... ha reaccionado ante el chasco[2] que se [3] llevó en ARCO
2. ... los aficionados al ciclismo [1] nos [3] llevaríamos una gran decepción [2]

F) LEXICALIZACIONES

1. Llevar una doble vida
2. ... lo llevarnos claro
3. ... lo llevar crudo

5 Trabajo en curso

En esta comunicación se ha presentado una metodología de análisis de los predicados en términos de su prototipicidad, con el fin de asignar usos y sentidos a niveles de aprendizaje en el marco de la adquisición de segundas lenguas. Se parte de una hipótesis inicial que equipara nivel de prototipicidad con nivel de aprendizaje, razón por la cual se han definido unos parámetros de prototipicidad y se han aplicado a algunos de los usos del verbo *llevar*.

Ahora bien, cabe destacar que la repercusión de cada uno de los factores en la determinación del grado de prototipicidad del verbo no se ha valorado con exactitud, pues es necesario realizar pruebas experimentales, de la misma manera que deben realizarse para comprobar la validez de la hipótesis de partida. A partir de este cuerpo teórico se desarrollará una batería de ejemplos experimentales. Mediante los experimentos se comprobará si la teoría es correcta.

Consideramos que esta investigación aportará nuevos datos, pues no se ha llevado a cabo hasta ahora un trabajo de modelización del conocimiento verbal por niveles de aprendizaje. El material didáctico existente sobre verbos, con excepción de trabajos parciales realizados en la Universidad de Utrecht bajo la dirección de P. Slagter, está enfocado a la enseñanza de la conjugación verbal en el sentido estrictamente paradigmático. Por esta razón consideramos que nuestra investigación puede ser útil, no sólo en el ámbito del aprendizaje de segundas lenguas, sino también en el ámbito de la lexicografía y en la validación del modelo teórico de que partimos.

Agradecimientos

Este trabajo ha sido elaborado gracias a las ayudas económicas del proyecto PB-94 0830 de la DGYCT, la acción APC-96 0125 de la DGICYT, ITEM TIC-96 1243-C03-02, y la beca predoctoral FI 97/00306 PG de la CIRIT de la Generalitat de Catalunya. Los datos del Español como L2 han sido recogidos gracias a un proyecto DGES PB-94-1096-C02-01 en coordinación con la UPF, UCM y U. de Ottawa.

Referencias bibliográficas

- [Bresnan & Kanerva 92] Bresnan, J. & J.M. Kanerva (1992) The Thematic Hierarchy and Locative Inversion in UG. a reply to Paul Schachter's comments. En *Syntax and Semantics 26.: Syntax and the Lexicon*, ed. E. Wehrli and T. Stowell, 111-125. San Diego, Calif.: Academic Press.
- [Coleman y Kay 81] Coleman, L. y P. Kay (1981) Prototype semantics. The English Word 'lie'. *Language* 57: 26-44.
- [Croft 90] Croft, W. (1990) Possible verbs and the structure of events. In: Tsohatzidis, S. L. Meanings and Prototypes. *Studies in Linguistic Categorization*. Routledge, London.
- [Fernández et al. 98] Fernández, A., Martí, M.A., Vázquez, G., Castellón, I. & R. Morante (1998). Criterios de clasificación de las estructuras eventivas: el cambio y la transferencia. *Actas del XIII Congreso de AESLA*. Universidad de la Rioja.
- [Kleiber 95] Kleiber, G. 1995. *La semántica de los prototipos*. Madrid: Visor Libros.
- [Kronenfeld 88] Kronenfeld, D. (1988) Full bloods and protestants: semantic extension in complex domains, Ms. University of California, Riverside.
- [Lakoff 87] Lakoff, G. (1987) *Women, fire, and dangerous things: what categories reveal about the mind*. University of Chicago Press.
- [Langacker 87] Langacker, R. (1987) *Foundations of cognitive grammar*. Stanford University Press.
- [Levin 93] Levin, B. (1993) *English Verb Classes and Alternations*. The University of Chicago Press.
- [Liceras 96] Liceras, J.M. (1996) *La adquisición de las lenguas segundas y la*

gramática universal. Madrid: Síntesis.

- [Little 98] Little, D. (1998) *Technologies, Media and Foreign Language Learning*. Authentik Language Learning Resources Ltd. Ireland.
- [Morante & Martí 98] Morante, R. & M.A. Martí (1998) Learning Catalan and Spanish Verbs through EuroWordNet. *Proceedings of the NLP+IA 98*. Moncton, NB, Canada.
- [Rosch 77] Rosch, E. (1977) Human categorization, en N. Warren, ed., *Advances in Cross-Cultural Psychology*, vol. I, Nueva York: Academic Press.
- [Self 92] Self, J. (1992) A framework for Learner Modelling. *Interactive Learning Environments*, Vol. 2, Issue (2), 111-137.
- [Talmy 85] Talmy, L. (1985) 'Lexicalization Patterns: semantic structure in lexical form' en T. Shopen (ed.) *Language typology and syntactic description*, vol. 3 CUP, Cambridge.
- [Taylor 89] Taylor, J.R. (1989) *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford, Clarendon Press.
- [Vázquez 98] Vázquez, G. (1998) *La interacció entre la sintaxi i la semàntica en el comportament verbal*. Doctoral Dissertation, Universitat de Lleida.
- [Vázquez et al. 99] Vázquez, G., Fernández, A., Martí, M.A., Morante, R. e I. Castellón (1999) A hierarchically structured LKB for verbs. *Proceedings of the Third International Workshop on Computational Semantics*. University of Tilburg.
- [Vossen 98] Vossen, P. (ed) (1998) *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*, Kluwer Academic Publishers, Dordrecht.

